

**Figure 1:** BLAST result against NCBI non-redundant database using SEQ ID NO: 6 (the INSP108 polypeptide).

BLASTP 2.2.2 [Jan-08-2002]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= INSP108.pp  
(77 letters)

Database: All non-redundant GenBank CDS  
translations+PDB+SwissProt+PIR+PRF  
1,267,376 sequences; 405,046,914 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
ref NP_697019.1  defensin, beta 123; defensin, beta 23 [Homo sap...	51	4e-06
ref XP_141520.1  similar to defensin, beta 123; defensin, beta 2...	46	1e-04
gb AAM93917.1  defensin beta 124 [Homo sapiens]	39	0.012
emb CAB72350.2  dJ1018D12.3 (a putative novel protein) [Homo sap...	39	0.016
ref NP_473453.1  epididymus specific clone 42; chromosome 20 ope...	39	0.016
ref NP_660139.1  defensin beta 119; testis-specific beta-defensi...	39	0.021
sp Q95LI0 D118_MACMU Beta-defensin 118 precursor (Epididymal sec...	37	0.079
ref NP_631968.1  defensin beta 15 [Mus musculus] >gi 19171622 em...	37	0.079
ref NP_689464.1  defensin, beta 106; defensin, beta 6 [Homo sapi...	35	0.18
gb AAN33114.1  beta-defensin 106 [Homo sapiens]	35	0.18

2/22

**Figure 2:** Alignment between INSP108 polypeptide sequence (SEQ ID NO:6) and defensin beta 123 (*Homo sapiens*).

```
>ref|NP_697019.1| defensin, beta 123; defensin, beta 23 [Homo sapiens]
sp|Q8N688|D123_HUMAN Beta-defensin 123 precursor (Beta-defensin 23) (DEFB-23)
gb|AAM93916.1| defensin beta 123 [Homo sapiens]
      Length = 67

Score = 50.8 bits (120), Expect = 4e-06
Identities = 26/67 (38%), Positives = 33/67 (48%), Gaps = 10/67 (14%)

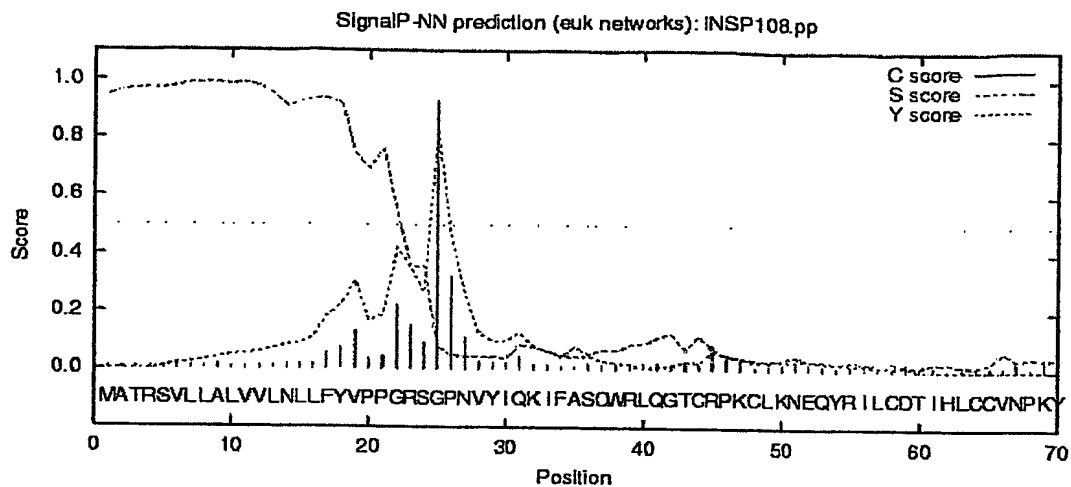
Query:  6  VLLALVVLNLLFYVPPGRSGPNVYIQIFASCWRLQGTCTRPKCLKNEQYRILCDTIHLCC 65
      +LL L VL LL  + PG +          CW L G CR +C K E+  + C    +CC
Sbjct:  4  LLLTLTVLLLLSQLTPGGT-----QRCWNLYGKCRYRCSKKERVYVYCINNKMCC 53

Query: 66  VNP KYLP 72
      V  PKY  P
Sbjct: 54  VKPKYQP 60
```

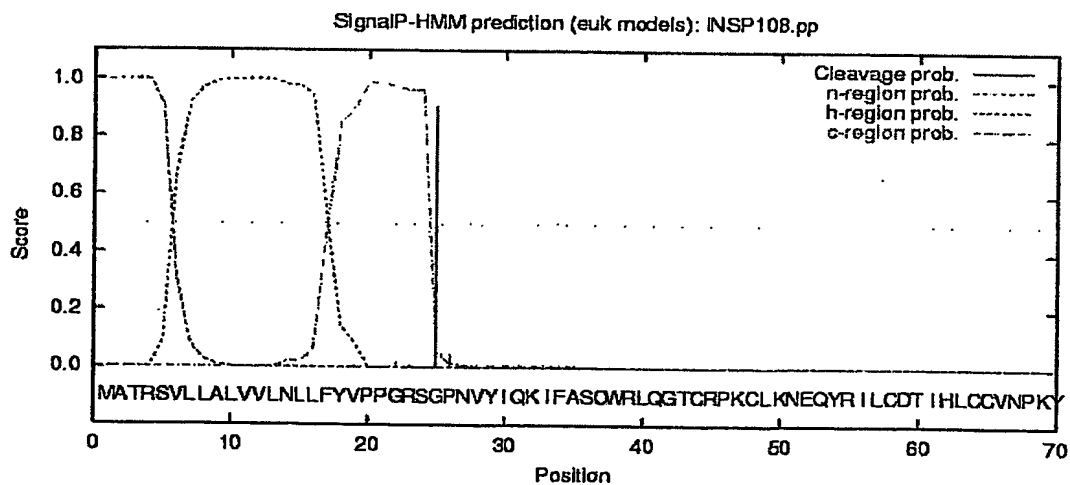
3/22

**Figure 3: Sig P cleavage site prediction for INSP108.**

&gt;INSP108.pp



```
>INSP108.pp                      length = 70
# Measure  Position  Value  Cutoff  signal peptide?
max. C      25      0.933  0.33   YES
max. Y      25      0.828  0.32   YES
max. S       9      0.991  0.82   YES
mean S     1-24      0.864  0.47   YES
# Most likely cleavage site between pos. 24 and 25: GRS-GP
```



```
>INSP108.pp
Prediction: Signal peptide
Signal peptide probability: 1.000
Signal anchor probability: 0.000
Max cleavage site probability: 0.906 between pos. 24 and 25
```

4/22

**Figure 4:** BLAST result against NCBI non-redundant database using SEQ ID NO: 14 (the INSP109 polypeptide).

BLASTP 2.2.2 [Jan-08-2002]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= INSP109.pp  
(78 letters)

Database: All non-redundant GenBank CDS  
translations+PDB+SwissProt+PIR+PRF  
1,267,376 sequences; 405,046,914 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
dbj BAC37510.1  unnamed protein product [Mus musculus]	88	3e-17
ref NP_062702.1  defensin beta 4; beta defensin-4 [Mus musculus]...	33	1.1
emb CAA08905.1  beta defensin-2 [Capra hircus]	31	3.3
ref NP_348802.1  Uncharacterized protein, homolog HI1244 from Ha...	31	4.3
gb AAG10514.1 AF288371_1 beta-defensin 4 variant [Mus musculus]	31	4.3
emb CAD23115.1  blue cone opsin [Cottus kesslerii]	30	7.4
ref NP_689464.1  defensin, beta 106; defensin, beta 6 [Homo sapi...	30	9.6
gb AAN33114.1  beta-defensin 106 [Homo sapiens]	30	9.6
ref XP_163302.1  hypothetical protein XP_163302 [Mus musculus]	30	9.6

5/22

**Figure 5:** Alignment between INSP109 polypeptide sequence (SEQ ID NO:14) and defensin beta 4 (*Mus musculus*).

```

>ref|NP_062702.1| defensin beta 4; beta defensin-4 [Mus musculus]
sp|P82019|BD04_MOUSE Beta-defensin 4 precursor (BD-4) (mBD-4)
gb|AAD38852.1|AF155882_1 beta defensin-4 [Mus musculus]
gb|AAG02197.1|AF287475_1 beta-defensin 4 precursor [Mus musculus]
dbj|BAB26051.1| unnamed protein product [Mus musculus]
dbj|BAB26207.1| unnamed protein product [Mus musculus]
      Length = 63

Score = 32.7 bits (73), Expect = 1.1
Identities = 19/56 (33%), Positives = 27/56 (47%), Gaps = 1/56 (1%)

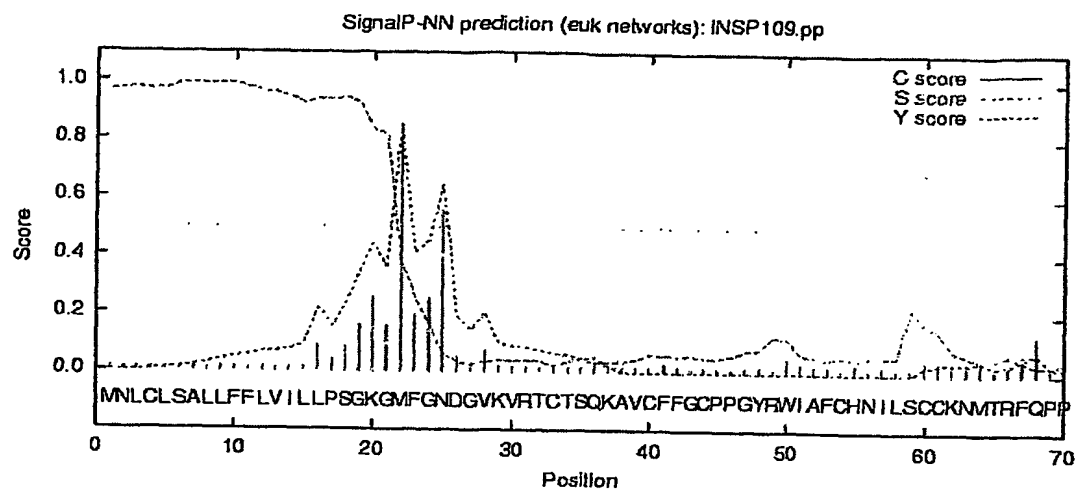
Query: 8  LLFFLVILLPSGKGMFGNDGVKVRTCTSQKAVCFFGCPPGYRWIAFC-HNILSCCK 62
          LLF  +++L S   F          TC + A+C+  CP +R I  C H  + CCK
Sbjct: 6  LLFTLLVLLSPLAAFTQIINNPICTMTNGAICWGPCPTAFRQIGNCGHFKVRCK 61

```

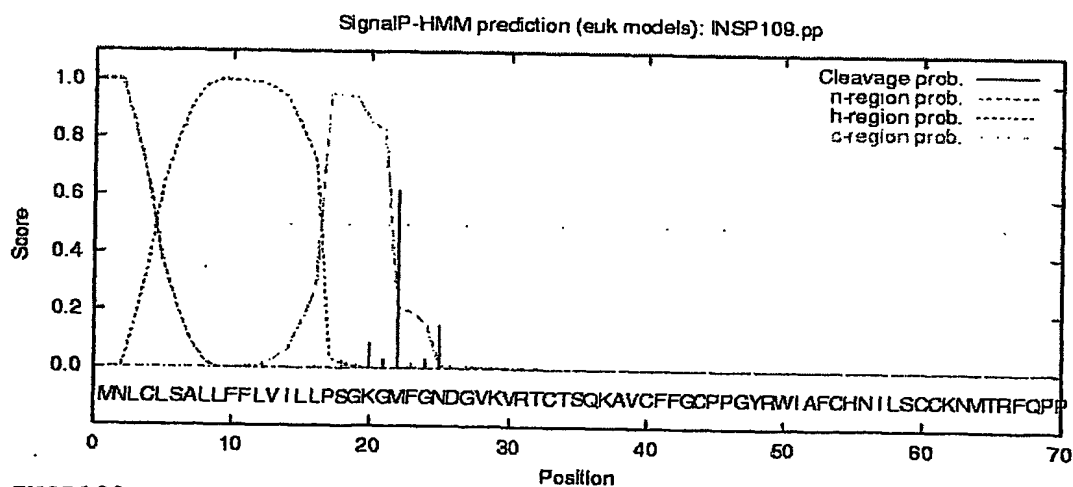
6/22

**Figure 6: Sig P cleavage site prediction for INSP109.**

&gt;INSP109.pp



```
>INSP109.pp                                length = 70
# Measure  Position  Value  Cutoff  signal peptide?
max. C      22       0.849  0.33   YES
max. Y      22       0.854  0.32   YES
max. S       7       0.994  0.82   YES
mean S     1-21      0.953  0.47   YES
# Most likely cleavage site between pos. 21 and 22: GKG-MF
```



```
>INSP109.pp
Prediction: Signal peptide
Signal peptide probability: 0.999
Signal anchor probability: 0.001
Max cleavage site probability: 0.619 between pos. 21 and 22
```

**Figure 7: Predicted nucleotide sequence of INSP108 with translation**

```

1  ttgttccaaa aggttcacta gccatgcagc tccccgtctc ttcaaagctg cggagagagt
61  gactctccga tgagtcacag ctgettcctt gctgattggt atggccacaa ggagcgtcct
                                     m a t r s v
                                     INSP108-F1
121  cttggccctc gtggtcctta acttactctt ctatgttcca ccaggtagaa gtggacccaa
    l l a l v v l n l l f y v p p g r s g p
181  tgtctacata caaaaaatct ttgcttcatg ttggcgactg caaggtactt gccggccaaa
    n v y i q k i f a s c w r l q g t c r p
241  atgtctaaaa aacgaacaat atcgtatttt gtgtgatact atacatttgt gctgtgtaaa
    k c l k n e q y r i l c d t i h l c c v
301  cccaaaatat ttacctatac tgactgggaa atagttgtga gtacctgaaa gctgttgctg
    n p k y l p i l t g k
                                     INSP108-R1
361  atttcctctg ggaaccaga tccctctcag ttgcaccatt cgattaaaac aatggcttta
421  gcctatcagt gttc

```

Position and sense of PCR primers →  
Underlined: signal sequence

**Figure 8:** Nucleotide sequence with translation of INSP108 PCR product cloned using primers INSP108-CP1 and INSP108-CP2.

```

1  attggtatgg ccacaaggag cgtcctcttg gccctcgtgg tccttaactt actcttctat
   m a t r s v l l a l v v l n l l f y
   INSP108-F1 →
61  gttccaccag gtagaagtgg acccaatgtc tacatacaaa aaatctttgc ttcattgttg
   v p p g r s g p n v y i q k i f a s c w
121 cgactgcaag gtacttgccg gccaaaatgt ctaaaaaacg aacaatatcg tattttgtgt
   r l q g t c r p k c l k n e q y r i l c
181 gatactatac atttgtgctg tgtaaaccga aaatatattac ctatactgac tgggaaatag
   d t i h l c c v n p k y l p i l t g k
241 ttgtg ← INSP108-R1

```

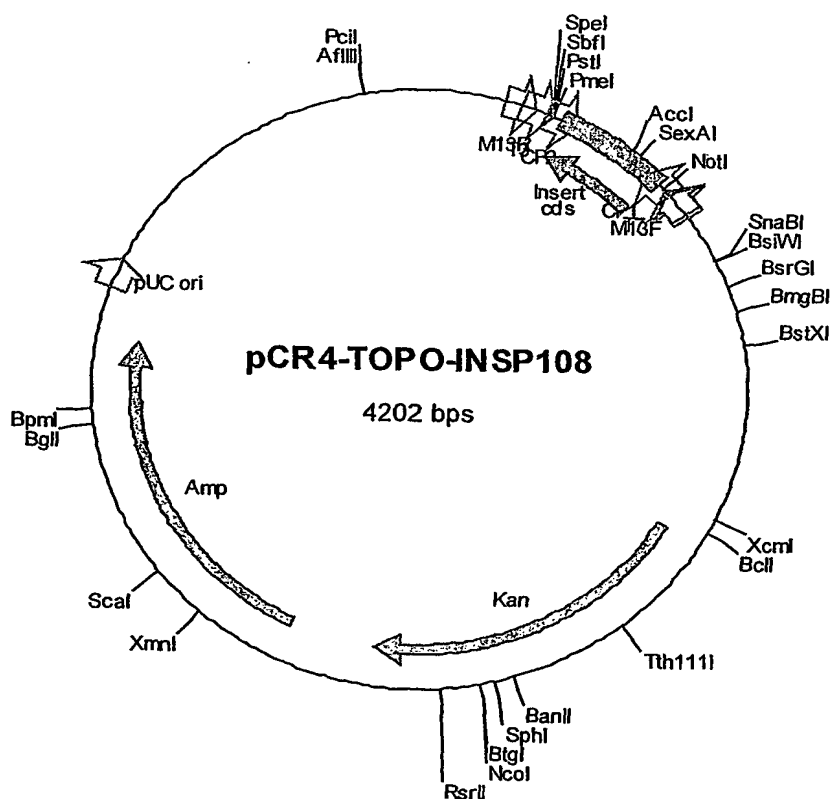
Position and sense of PCR primers →



**Figure 9: Map of pCR4-TOPO-INSP108**

Molecule: pCR4-TOPO-INSP108, 4202 bps DNA Circular

Type	Start	End	Name	Description
MARKER	205		M13R	M13 rev priming site
MARKER	243		T3	T3 priming site
MARKER	295		CP2	INSP108-CP2
GENE	533	303	C cds	INSP108 cds
MARKER	539		C CP1	INSP108-CP1
REGION	539	295	C Insert	INSP108-F1R1 PCR product
MARKER	592		C T7	T7 priming site
MARKER	600		C M13F	M13 for priming site
GENE	1404	2198	Kan	Kanamycin resistance gene ORF
GENE	2402	3262	Amp	Ampicillin resistance gene ORF
MARKER	3407		pUC ori	pUC origin

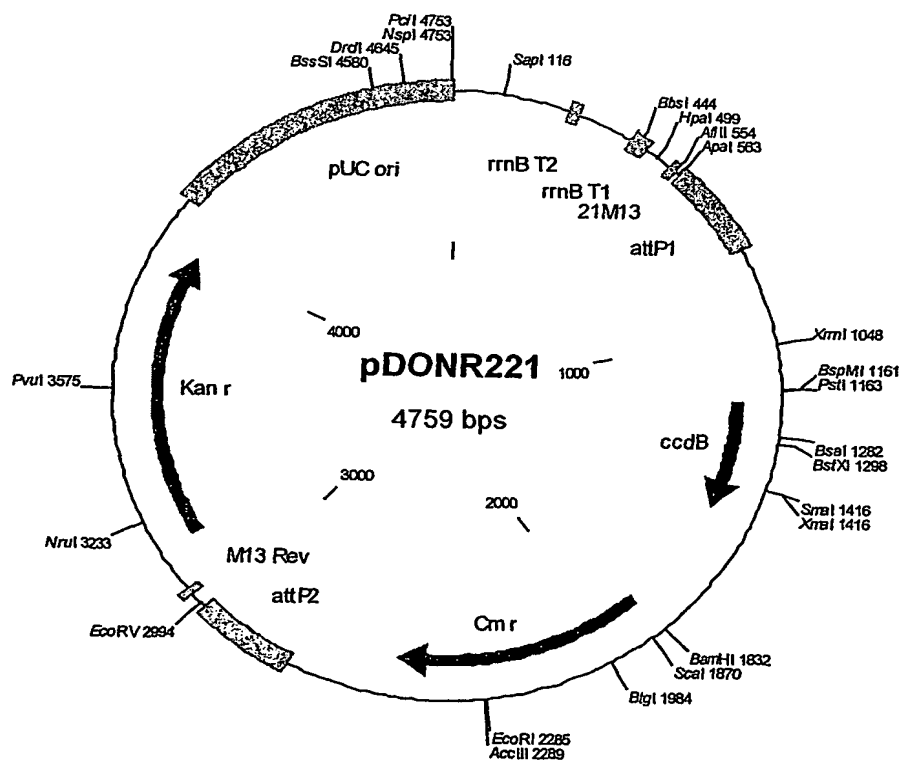


**Figure 10: Map of pDONR 221**

Molecule: pDONR221, 4759 bps DNA Circular  
 File Name: pDONR221.cm5

**Description:**

Type	Start	End	Name	Description
REGION	295	268	C rrnB T2	transcription termination sequence
REGION	470	427	C rrnB T1	transcription termination sequence
REGION	536	553	21M13	M13 Forward primer
REGION	570	801	attP1	
GENE	1197	1502	ccdB	
GENE	1844	2503	Cm r	Chloramphenicol resistance gene
REGION	2751	2982	attP2	
REGION	3040	3023	C M13 Rev	M13 Reverse primer
GENE	3153	3962	Kan r	
REGION	4083	4756	pUC ori	



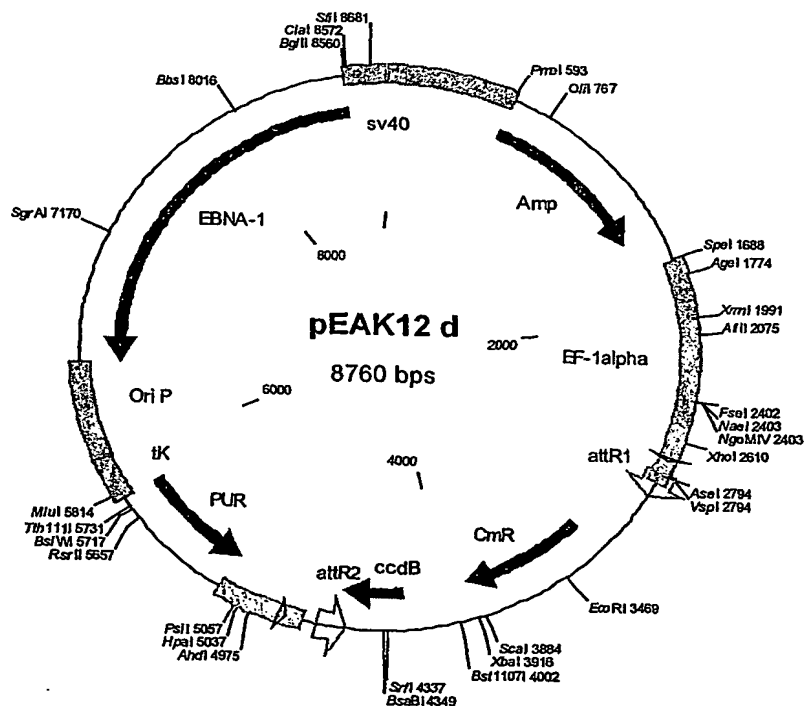
**Figure 11: Map of expression vector pEAK12d**

Molecule: pEAK12 d, 8760 bps DNA Circular  
 File Name: pEAK12DEST.cm5

Description: Mammalian cell expression vector (plasmid ID 11345)

## Molecule Features:

Type	Start	End	Name	Description
REGION	2	595		pmb-ori
GENE	596	1519	Amp	
REGION	1690	2795	EF-1alpha	
REGION	2703	2722		position of pEAK12F primer
REGION	2796	2845		MCS
MARKER	2855		attR1	
GENE	3256	3915	CmR	
GENE	4257	4562	ccdB	
MARKER	4603		C attR2	
REGION	4733	4733		MCS
REGION	4734	5162		poly A/splice
REGION	4819	4848	C	position of pEAK12R primer
GENE	5781	5163	C PUR	PUROMYCIN
REGION	6005	5782	C tK	tK promoter
REGION	6500	6006	C Ori P	
GENE	8552	6500	C EBNA-1	
REGION	8553	8752	sv40	

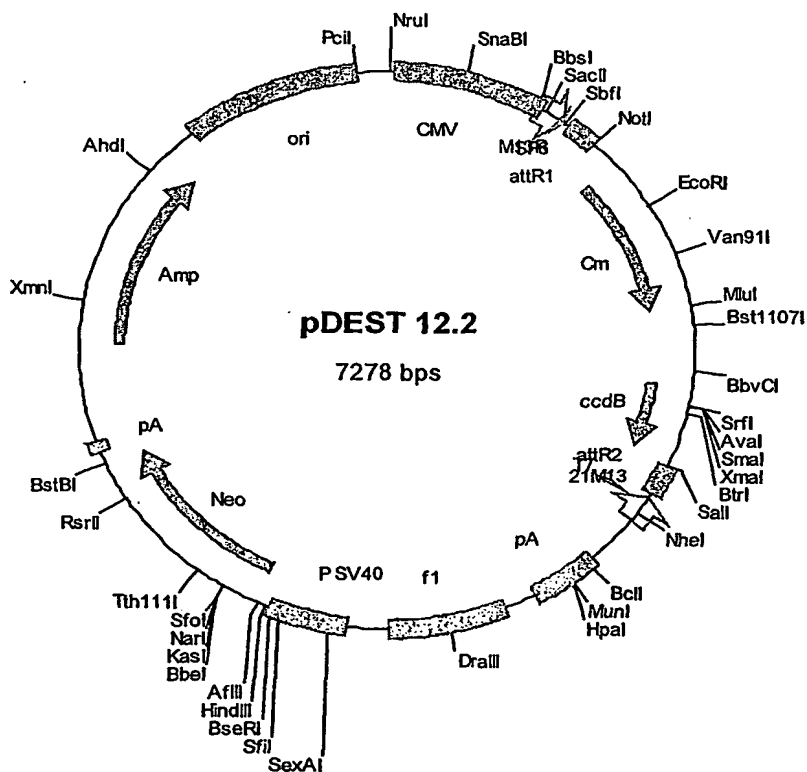


**Figure 12: Map of expression vector pDEST12.2**

Molecule: pDEST 12.2, 7278 bps DNA Circular  
 File Name: pDEST12-2.cm5

Description: Eukaryotic expression vector

Type	Start	End	Name	Description
REGION	15	608	CMV	CMV promoter
MARKER	648		M13R	M13R primer
REGION	687	706	SP6	SP6 promoter
REGION	730	854	attR1	
GENE	963	1622	Cm	
GENE	1964	2269	ccdB	
REGION	2310	2434	attR2	
GENE	2484	2464	C T7	T7 promoter
MARKER	2512		C 21M13	21M13 primer
REGION	2784	3050	pA	SV40 polyadenylation signal
REGION	3176	3631	f1	f1 intergenic region
REGION	3791	4099	P SV40	SV40 ori & early promoter
GENE	4158	4952	Neo	
REGION	5016	5064	pA	synthetic poly adenylation signal
GENE	5475	6335	Amp	
REGION	6480	7153	ori	pUC ori

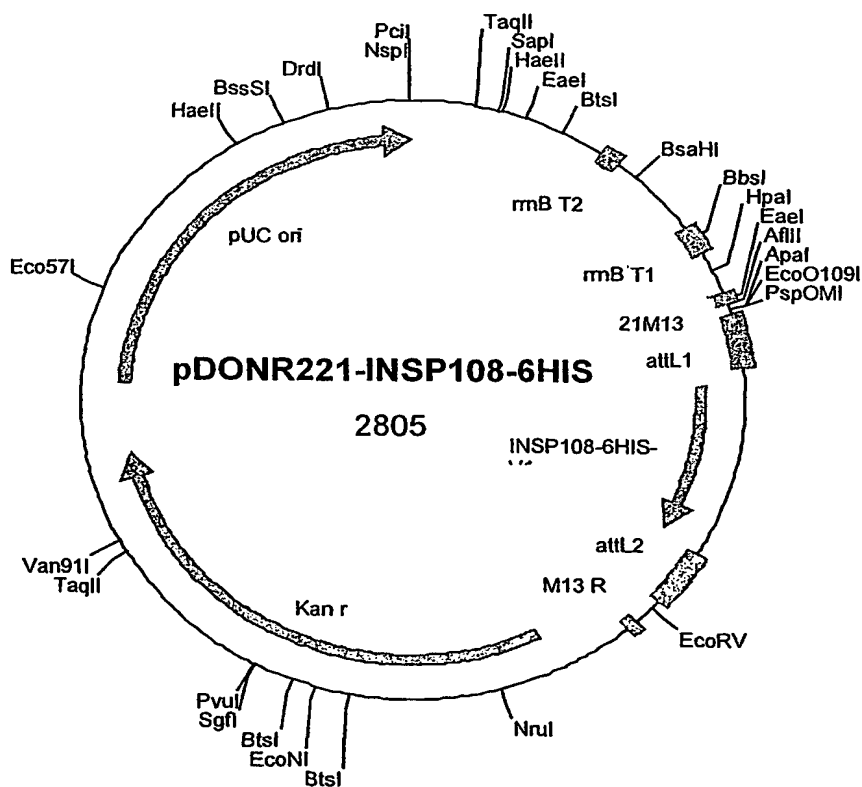


13/22

**Figure 13: Map of pDONR221-INSP108-6HIS**

Molecule: pDONR221-INSP108-6HIS, 2805 bps DNA Circular

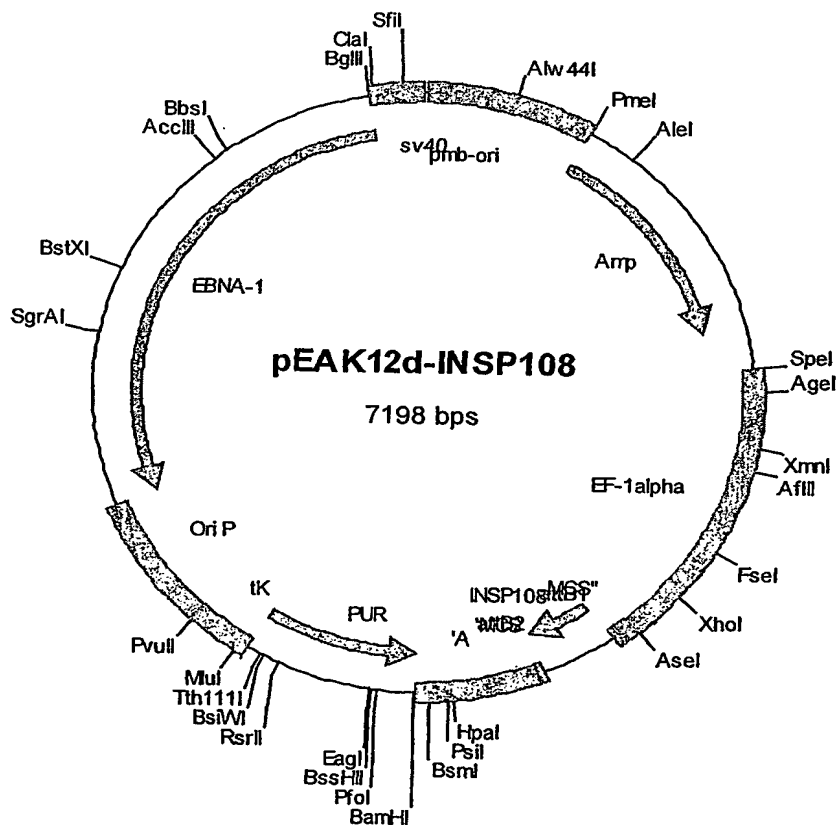
Type	Start	End	Name	Description
REGION	295	268	C rrnB T2	transcription termination sequence
REGION	470	427	C rrnB T1	transcription termination sequence
REGION	536	553	21M13	21M13 primer
REGION	570	651	attL1	
GENE	677	925	INSP108-6HIS-V1	INSP108-6HIS ORF
REGION	940	1028	attL2	
REGION	1086	1070	C M13 R	M13R primer
GENE	1199	2008	Kan r	
GENE	2129	2802	pUC ori	



**Figure 14: Map of pEAK12d-INSP108-6HIS**

Molecule: pEAK12d-INSP108, 7198 bps DNA Circular

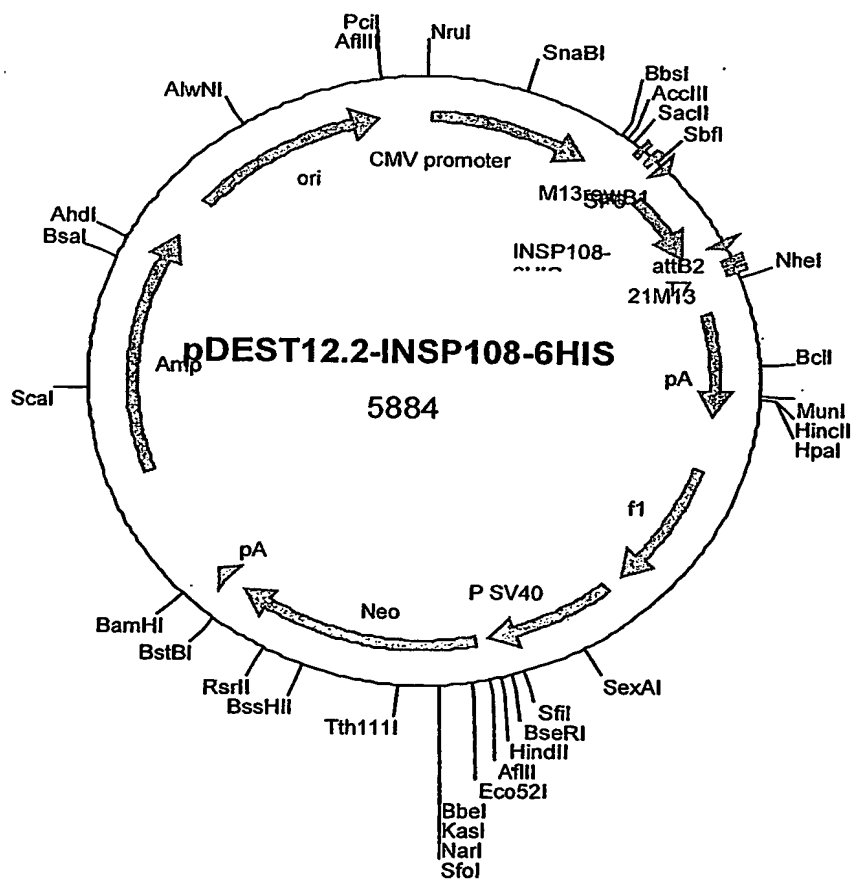
Type	Start	End	Name	Description
REGION	2	595	pmb-ori	
GENE	596	1519	Amp	
REGION	1690	2795	EF-1alpha	
REGION	2796	2845	MCS''	
REGION	2855	2874	attB1	
GENE	2888	3139	INSP108	
REGION	3144	3165	attB2	
REGION	3171	3171	'MCS	
REGION	3172	3600	'A	poly A/splice
GENE	4219	3601 C	PUR	PUROMYCIN
REGION	4443	4220 C	tK	tK promoter
REGION	4938	4444 C	Ori P	
GENE	6990	4938 C	EBNA-1	
REGION	6991	7190	sv40	



**Figure 15: Map of pDEST12.2-INSP108-6HIS**

Molecule: pDEST12.2-INSP108-6HIS, 5884 bps DNA Circular

Type	Start	End	Name	Description
GENE	15	537	CMV promoter	
REGION	648	665	M13rev	M13R primer
REGION	687	704	SP6	SP6 primer
REGION	730	762	attB1	
GENE	763	1011	INSP108-6HIS	
REGION	1015	1040	attB2	
REGION	1090	1070	C T7	T7 promoter
REGION	1119	1101	C 21M13	21M13 primer
GENE	1225	1587	pA	SV40 polyadenylation signal
GENE	1781	2237	f1	f1 intergenic region
GENE	2301	2719	P SV40	SV40 ori & early promoter
GENE	2764	3558	Neo	
GENE	3622	3670	pA	poly adenylation signal
GENE	4081	4941	Amp	
GENE	5090	5729	ori	pUC ori



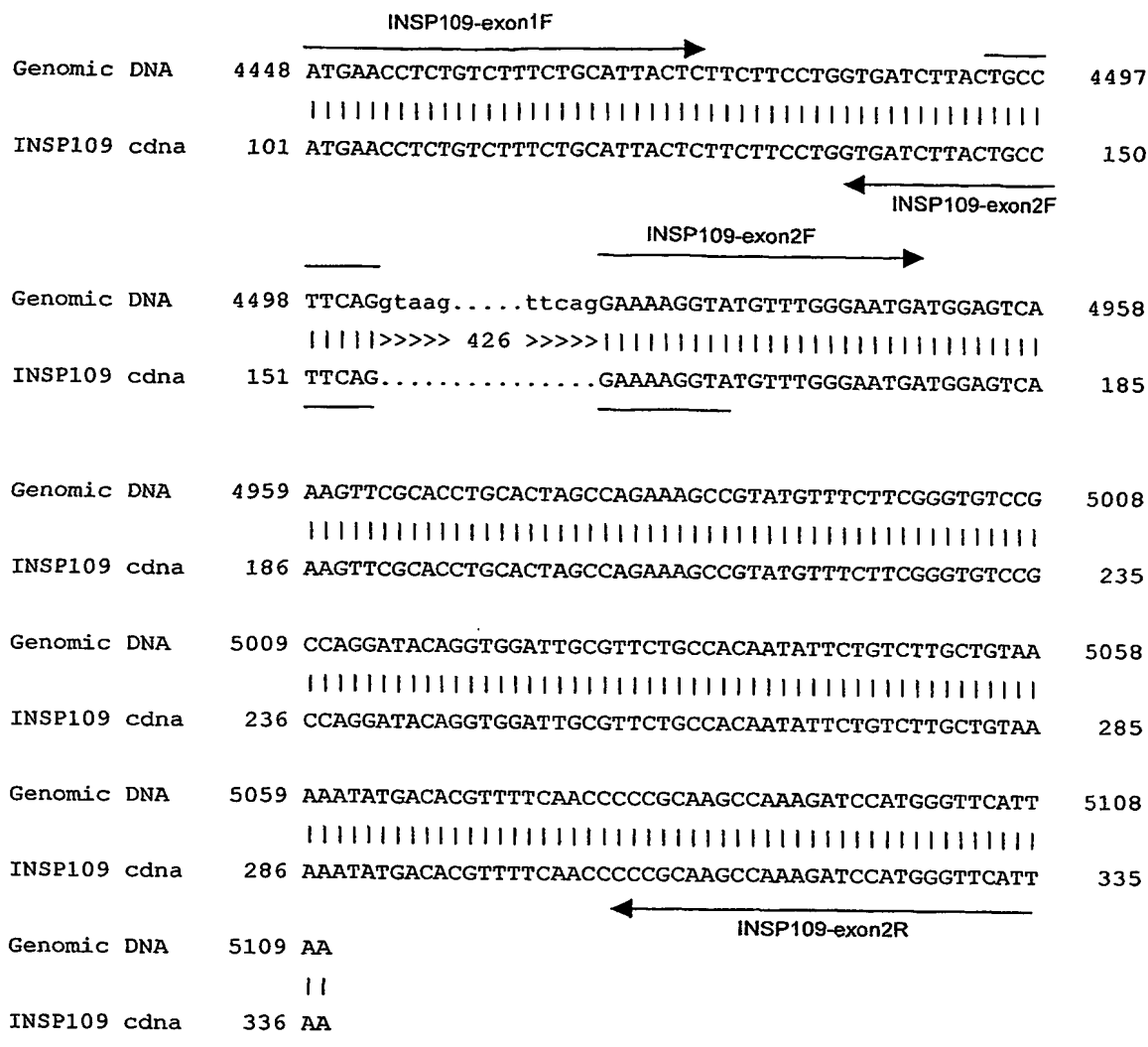
16/22

**Figure 16:** Predicted nucleotide sequence of INSP109 with translation

```
1  atgaacctct gtctttctgc attactcttc ttcctggtga tcttactgcc ttcaggaaaa
   m n l c l s a l l f f l v i l l p s g k
61  ggtatgtttg ggaatgatgg agtcaaagtt cgcacctgca ctagccagaa agccgtatgt
   g m f g n d g v k v r t c t s q k a v c
121 ttcttcgggt gtccgccagg atacaggtgg attgcgttct gccacaatat tctgtcttgc
    f f g c p p g y r w i a f c h n i l s c
181 tgtaaaaata tgacacgttt tcaacccccg caagccaaag atccatgggt tcat
    c k n m t r f q p p q a k d p w v h
```

Underlined= signal peptide



**Figure 17:** INSP109 coding exon organization in genomic DNA and position of PCR primers

**Figure 18:** Nucleotide sequence and translation of cloned INSP109 ORF

```
1  atgaacctct gtctttctgc attactcttc ttcttggtga tcttactgcc ttcaggaaaa
   m n l c l s a l l f f l v i l l p s g k

61  ggtatgtttg ggaatgatgg agtcaaagtt cgcacctgca ctagccagaa agcogtatgt
   g m f g n d g v k v r t c t s q k a v c

121 ttcttcgggt gtccgccagg atacaggtgg attgcgttct gccacaatat tctgtcttgc
   f f g c p p g y r w i a f c h n i l s c

181 tgtaaaaata tgacacgttt tcaacccccg caagccaaag atccatgggt tcatta
   c k n m t r f q p p q a k d p w v h
```

**Figure 19: Map of pCR4-TOPO-INSP109**

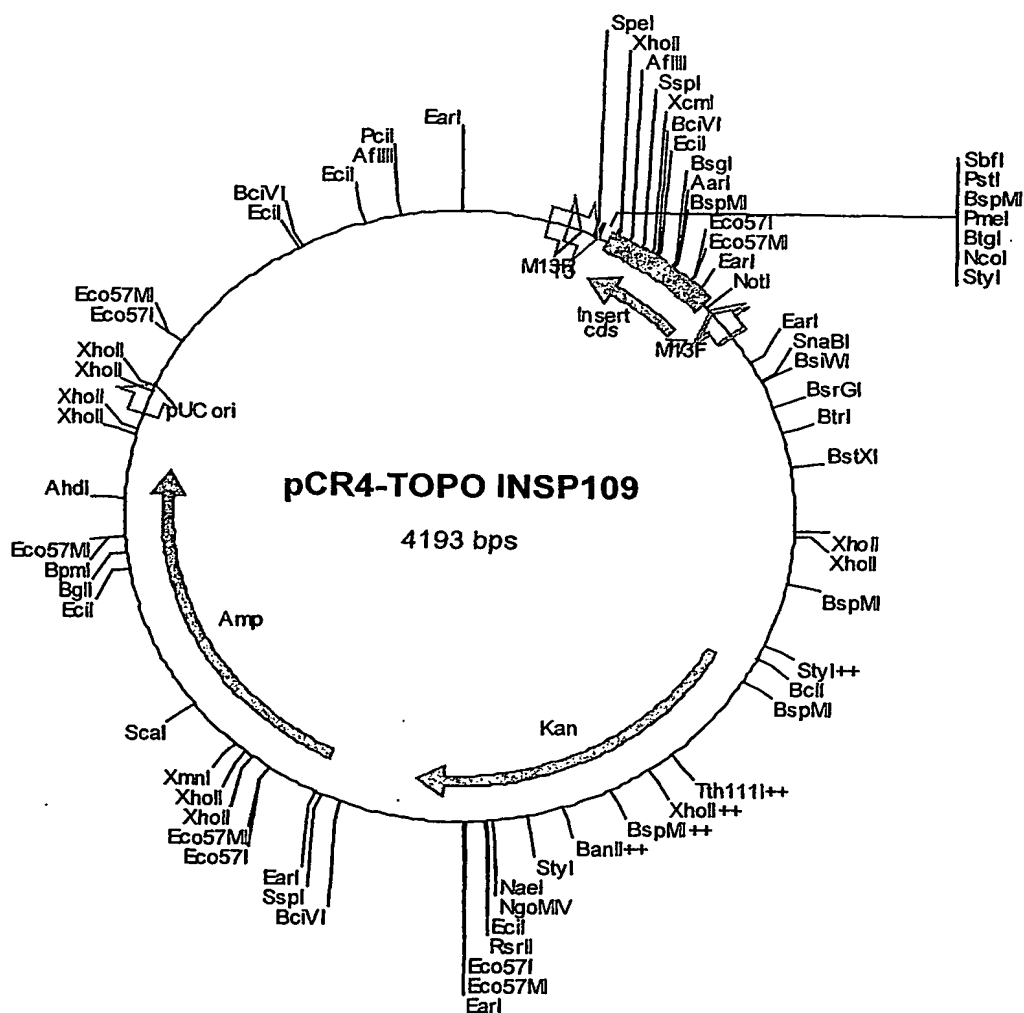
25 Nov 2003

## Molecule Features

Molecule: pCR4-TOPO INSP109, 4193 bps DNA Circular  
 File Name: 13984[1].cm5

Description: Ligation of inverted INSP109 assembled insert into pCR4-TOPO linear vector\*

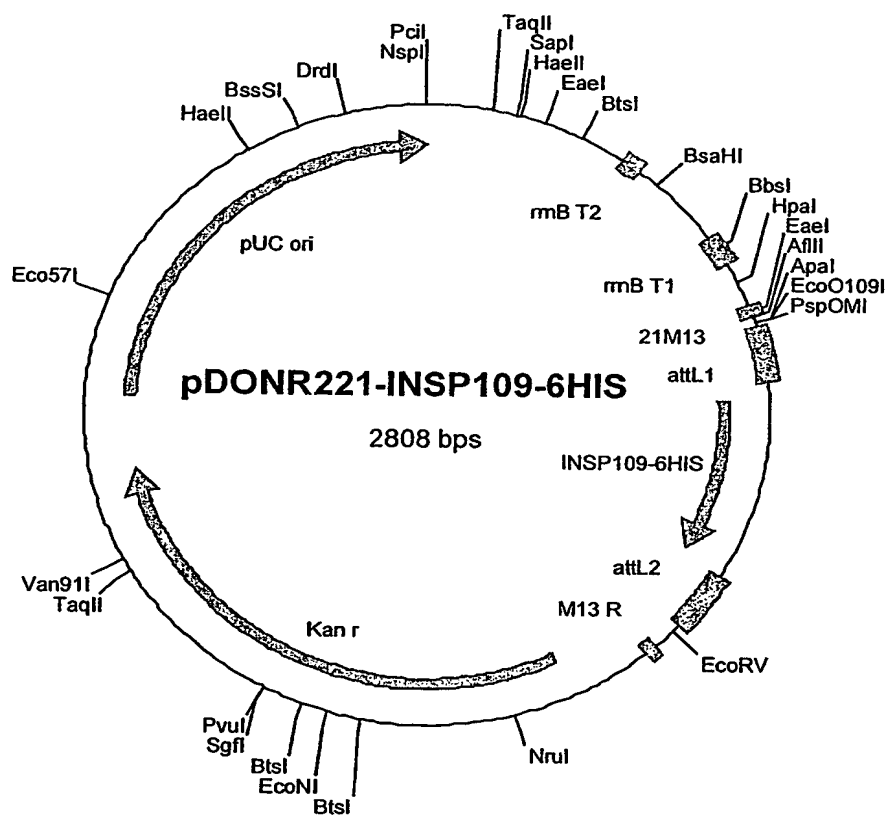
Type	Start	End	Name	Description
MARKER	205		M13R	M13 rev priming site
MARKER	243		T3	T3 priming site
REGION	530	295	C Insert	INSP109 assembled insert
GENE	530	297	C cds	INSP109 cds
MARKER	583		C T7	T7 priming site
MARKER	591		C M13F	M13 for priming site
GENE	1395	2189	Kan	Kanamycin resistance gene ORF
GENE	2393	3253	Amp	Ampicillin resistance gene ORF
MARKER	3398		pUC ori	pUC origin



**Figure 20: Map of pDONR-INSP109-6HIS**

Molecule: pDONR221-INSP109-6HIS, 2808 bps DNA Circular

Type	Start	End	Name	Description
REGION	295	268	C rrnB T2	transcription termination sequence
REGION	470	427	C rrnB T1	transcription termination sequence
REGION	536	553	21M13	21M13 primer
REGION	570	651	attL1	
GENE	677	928	INSP109-6HIS	
REGION	943	1031	attL2	
REGION	1089	1073	C M13 R	M13R primer
GENE	1202	2011	Kan r	
GENE	2132	2805	pUC ori	

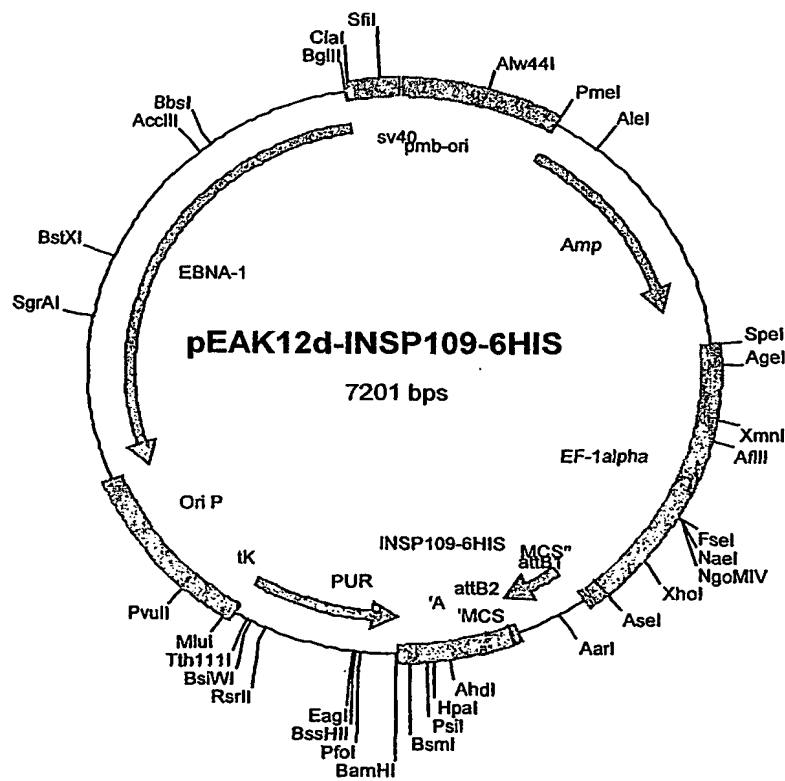


21/22

**Figure 21: Map of pEAK12d-INSP109-6HIS**

Molecule: pEAK12d-INSP109-6HIS, 7201 bps DNA Circular

Type	Start	End	Name	Description
REGION	2	595	pmb-ori	
GENE	596	1519	Amp	
REGION	1690	2795	EF-1alpha	
REGION	2796	2845	MCS''	
REGION	2855	2874	attB1	
GENE	2888	3142	INSP109-6HIS	
REGION	3147	3168	attB2	
REGION	3174	3174	'MCS	
REGION	3175	3603	'A	poly A/splice
GENE	4222	3604 C	PUR	PUROMYCIN
REGION	4446	4223 C	tK	tK promoter
REGION	4941	4447 C	Ori P	
GENE	6993	4941 C	EBNA-1	
REGION	6994	7193	sv40	



**Figure 22: Map of pDEST12.2-INSP109-6HIS**

Molecule: pDEST12.2-INSP109-6HIS, 5887 bps DNA Circular

Type	Start	End	Name	Description
GENE	15	537	CMV promoter	
REGION	648	665	M13rev	M13R primer
REGION	687	704	SP6	SP6 primer
REGION	730	762	attB1	
GENE	763	1014	INSP109-6HIS	
REGION	1018	1043	attB2	
REGION	1093	1073 C	T7	T7 promoter
REGION	1122	1104 C	21M13	21M13 primer
GENE	1228	1590	pA	SV40 polyadenylation signal
GENE	1784	2240	f1	f1 intergenic region
GENE	2304	2722	P SV40	SV40 ori & early promoter
GENE	2767	3561	Neo	
GENE	3625	3673	pA	poly adenylation signal
GENE	4084	4944	Amp	
GENE	5093	5732	ori	pUC ori

